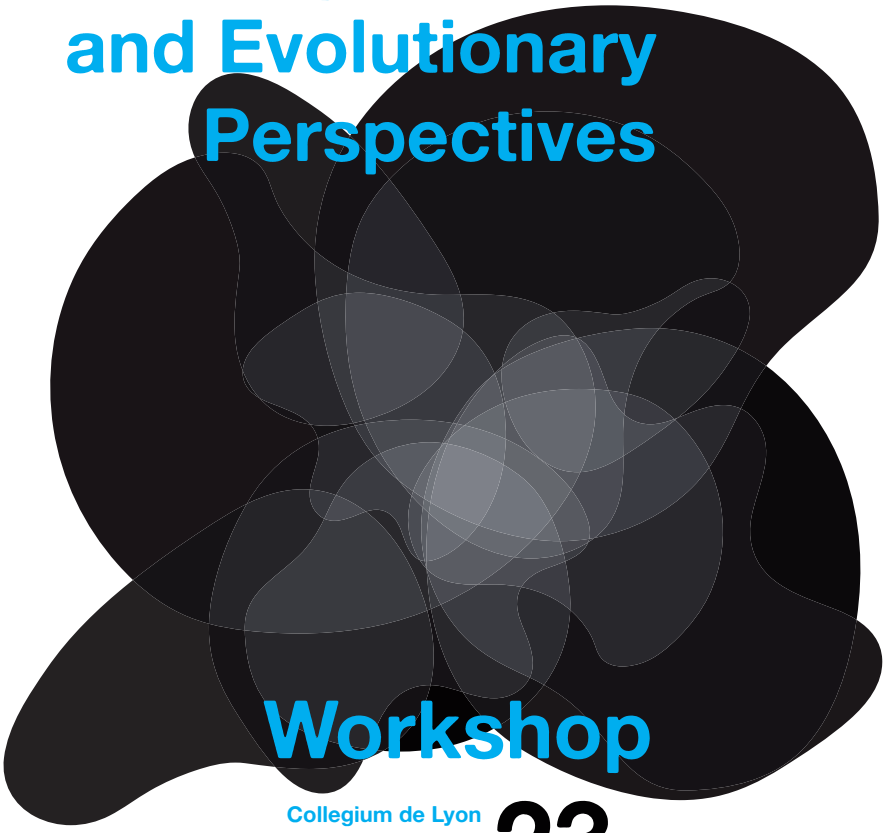
Complexity in Language:

Developmental and Evolutionary Perspectives



Workshop

Collegium de Lyon
ENS de Lyon

Room F 106

23-24 May 2011

PROGRAM

MAY 23-24

PROGRAM

Workshop on **Room F 106**
"Complexity in Language: Developmental and Evolutionary Perspectives"

Linguists have become increasingly interested in the subject matter of complexity. However, it's not clear that they have stopped to articulate explicitly what complexity means and what yardstick can be used to measure it, especially in studies comparing different language varieties. There has also been little discourse on alternative ways of conceiving of complexity and on what the different conceptions entail for accounts of the architectures and functions of languages. Can linguistics take inspiration from complexity theory as developed in, for instance, cybernetics and physics? This workshop is intended to foster a dialog between scholars of diverse but complementary backgrounds on these topics, to which we add developmental and evolutionary perspectives. How does complexity in language emerge ontogenetically and how did it evolve phylogenetically? Are there really languages that are less complex than others, in the same ways that successive stages of language development can display different levels of complexity? Is there a static kind of linguistic complexity that is distinct from dynamic complexity arising from the interactions of various modules during communication?

MAY 23 MONDAY

9:00	Welcome
	Session chair: Jean-Marie Hombert
9:30	**William S-Y. Wang**, Chinese University of Hong Kong / "Language and Complex Adaptive Systems."
10:30	**William Croft**, University of New Mexico / "Language complexity really lies in social cognition."
11:30	**Tom Schoenemann**, Indiana University / "A complex adaptive systems approach to language and brain"
12:30 - 2:00 PM	Lunch break
	Session chair: Didier Demolin
2:00	**Barbara Davis**, University of Texas at Austin & Collegium de Lyon /"Emergence of Complexity: The Case of Phonological Acquisition."
3:00	**Lucía Loureiro Porto**, University of Palma de Majorca / "The complexity of social interactions in language competition: Agent-based models and complex networks."
4:00 – 4:30	Coffee break
4:30	**Session chair: Salikoko S. Mufwene** **Albert Bastardas i Boada**, University of Barcelona / "Complexity in/of language: Bases for a dynamic sociocognitive view."
5:30	**Fermin Moscoso del Prado**, Dynamique du Langage, Université Lyon 2 / "Linguistic complexity as optimal grammar length: An information-theoretic approach."
6:30	Recess

MAY 24 TUESDAY

8:30	Welcome
	Session chair: François Pellegrino
9:00	**Jean-Marie Hombert**, Dynamique du Langage, Université Lyon 2 / "Population size, social complexity, and language dispersal in Bantu."
10:00	**Salikoko S. Mufwene**, University of Chicago & Collegium de Lyon / "The Emergence of Complexity in Language: An Evolutionary Perspective."
11:00	**Luc Steels**, University of Brussels (VUB-AI Lab) and Sony CSL, Paris / "Explaining the Origins of Complexity in Language: A case study for agreement systems."
12:00 - 1:30 PM	Lunch break
	Session chair: Christophe Coupé
1:30	**Vittorio Loreto**, Sapienza University of Rome and Institute for Scientific Interchange, Torino / "Statistical physics of language dynamics."
2:30	**Ramon Ferrer i Cancho**, Universitat Politècnica de Catalunya / "Word order as a constraint satisfaction problem: A mathematical approach."
3:30 – 4:00	Coffee break
	Session chair: Fermin Moscoso del Prado
4:00	**Christophe Coupé, Egidio Marsico, & François Pellegrino,** Dynamique du Langage,Université Lyon 2 "Complexity of phonological systems: Synchronic analyses & evolutionary models"
5:00	**Bart de Boer**, University of Amsterdam / "Self-organization in language and its relation to complexity."
6:00	End/Closing
7:00	Workshop Dinner

Collegium de Lyon / ENS de Lyon
Room F 106

ALBERT BASTARDAS I BOADA / BART DE BOER
RAMON FERRER I CANCHO / CHRISTOPHE COUPÉ
EGIDIO MARSICO / FRANÇOIS PELLEGRINO
WILLIAM CROFT / BARBARA DAVIS
JEAN-MARIE HOMBERT / VITTORIO LORETO
LUCÍA LOUREIRO PORTO / SALIKOKO S. MUFWENE
FERMIN MOSCOSO DEL PRADO / TOM SCHOENEMANN
LUC STEELS / WILLIAM S-Y. WANG

ABSTRACTS

Workshop on
COMPLEXITY IN LANGUAGE: DEVELOPMENTAL
AND EVOLUTIONARY PERSPECTIVES

MAY 23-24
2011

2

Albert Bastardas i Boada
University of Barcelona

"Complexity in/of language: Bases for a dynamic sociocognitive view."

There is no clear and accepted general definition for the term *complexity*. Several authors and research clusters use this word, but having different conceptual meanings in mind. In Linguistics, some approaches mention 'complex systems' in an attempt to put together new formal models, and thus try to apply this concept to different phenomena which show self-organization and emergence properties. There is also a philosophical-anthropological general 'complexity' thinking which is especially inspiring to sociocultural sciences. My work has been mainly based on the last approach.
I have been trying to build a holistic/autonomous and dynamic perspective in order to understand language contact and conflict evolution. This complex multidimensional approach conceives an ecosystem of language which includes brain/mind, social interaction, group, economic, media, and political dimensions. All of these dynamically interpenetrate each other, and, in doing so, co-produce and co-determine language forms, uses, and evolutions.
A transdisciplinary 'sociocomplexity' perspective based on human co-evolution and hologramic conceptualization, could combine both the computational and philosophical approaches, and be very useful to build a more complete understanding of language, mind, and society.

Bart de Boer
University of Amsterdam

"Self-organization in language and its relation to complexity."

Complexity in language is not necessarily associated with cognitive complexity.
In this talk I show how complexity can emerge through self-organization caused by interactions between language users in a population. This will be illustrated with two examples: that of vowel systems and that of combinatorial phonemic structure. It will be argued that apparently complex patterns can emerge without the language users being aware of them. This defines the difference between superficial structure and productive structure (where users are aware of the structure). Modern humans probably use most complex structures of language productively, but the possibility of complex structure being present superficially but not being used productively does provide a potential evolutionary pathway from simpler communication systems to complex linguistic ones. I will also present a novel experimental technique that can help distinguish between superficial and productive use of structure in systems of signals.

4

Ramon Ferrer i Cancho
Universitat Politècnica de Catalunya

“Word order as a constraint satisfaction problem: A mathematical approach.”

Consider the problem of arranging sequentially a head and its modifiers. We study the optimal placement of a head under two constraints: (a) maximizing the predictability of the sequence and (b) minimizing the online memory needed for handling the head-modifier dependencies as the sequence is produced. We show various mathematical results that illuminate various empirical findings on order « preferences ».

William Croft
University of New Mexico

"Language complexity really lies in social cognition."

Research in the evolution of linguistic complexity in the origin of language frequently focuses on the emergence of syntactic complexity: from one-word to multiword utterances, the emergence of hierarchical and recursive syntactic structures, or the differentiation of syntactic categories. I suggest that these syntactic processes, which are not insignificant, are less important for understanding the emergence of language than the evolution of social cognition (e.g. Clark 1996; Tomasello 2008) and conceptual processes in verbalization (Chafe 1977; Croft 2007). A framework of the basics of social cognition and verbalization will be presented, and some speculations may be offered regarding an evolutionary sequence for these social cognitive and conceptual abilities.

Christophe Coupé, Egidio Marsico, & François Pellegrino
Dynamique du Langage, Université Lyon 2

"Complexity of phonological systems: Synchronic analyses & evolutionary models"

The sound systems of the world languages exhibit properties that fit the framework of complex adaptive systems well.
Their multi-layered structure - sound systems composed of segments (vowels, consonants or diphthongs) which can themselves be decomposed into smaller units (features) – leads to a huge space of possible types. Yet, observed systems, while being highly variable, display significant regularities in the distributions and interactions of segments and features. Explanations for this organization have to be found in the individual histories of languages as well as in general constraints, whether articulatory, perceptive, or cognitive. In the last decades, various theoretical frameworks have been put forward to describe the latter mechanisms, focusing on concepts such as 'feature economy', 'ease of articulation', 'ease of perception', 'sufficient' or 'maximal' dispersion, etc.
Previous approaches are often tested against a variety of languages. Sometimes the differences remain more qualitative than quantitative. Additionally, if many studies have focused on vocalic systems, consonantal systems remain more elusive. We therefore try to offer a range of quantitative approaches

that depart from more traditional linguistic approaches. To this end, we have been 'data-mining' the UPSID database, which contains a carefully balanced sample of 451 languages from all linguistic families. We will review a range of attempts at deciphering the complexity of sound systems, taking into account the limits of our dataset:

i) revisiting the notion of feature economy with respect to the description length of the linguistic descriptions,
ii) analyzing the boundedness of the dataset and its statistical consequences,
iii) trying to derive evolutionary models from synchronic data.

Barbara Davis
University of Texas at Austin
& Collegium de Lyon

"Emergence of Complexity:
The Case of Phonological Acquisition."

Acquisition of the phonological component of language presents an opportunity to view emergence of a dynamically changing and complex system. In an emergence view, convergence of phonological knowledge and associated behavioral properties in human children is based on supporting mechanisms, including self-organization and learning. These mechanisms are domain-general and functional in their implementation across varied operations of the child within the environment. Phonological emergence is also conceptually reliant for expression on child social interaction capacities and input from environmental models. Importantly, phonological emergence is not based uniquely on expression through maturation of a modular and a priori linguistic competence/performance system as proposed in underlying grammar (UG) approaches. Operation of the complex system supporting emergence of phonology will be illustrated by cross-language corpora on early periods of phonological acquisition in children within typologically diverse language environments.

Jean-Marie Hombert
Dynamique du Langage, Université Lyon 2

“Population size, social complexity, and language dispersal in Bantu.”

Linguistic complexity can be affected by language contacts. In such cases the number of speakers of each linguistic community and the “prestige” of the languages in contact can play a determinant role in the evolution of the linguistic situation. In this paper we will present a case study of linguistic contacts between Bantu speakers expanding over most of sub-equatorial Africa over the last four millennia and hunter-gatherers (“Pygmies” and Khoesan). It is generally assumed that the “success” of this Bantu expansion is linked to a demographic superiority of Bantu speakers resulting from plant domestication. Increase in population size is often associated with an increase in social complexity. We will examine the interactions between these different factors and draw a chronological scenario for these contacts and their results on the linguistic systems of these populations.

Vittorio Loreto
Sapienza University of Rome
and Institute for Scientific Interchange, Torino

"Statistical physics of language dynamics."

Language dynamics is an emerging field that focuses on all processes related to the emergence,evolution, and extinction of languages. Recently the study of the self-organization and evolution of language and meaning has led to the idea that a community of language users can be seen as a complex dynamical system that collectively solves the problem of developing a shared communication framework through the back-and-forth signaling between people. In this talk I'll review some of the progresses made in the last few years and highlight potential future directions for the research in this area. I'll discuss in particular several examples corresponding to the early stages of the emergence of a language, namely the emergence of a common lexicon and the emergence of a shared set of linguistics categories. I'll point out how synthetic modeling has nowadays reached sufficient maturity to contribute significantly to the ongoing debate in cognitive science. For instance it has been recently possible to reproduce in a numerical model the outcomes of an important experimental survey, the so-called World Color Survey (WCS). In addition, new experimental frameworks are becoming progressively available. Finally I'll discuss the crucial issue in linguistics of whether structures

of languages we adopt are the outcome of an individual-based process of optimization or rather the result of a complex socially-driven cultural negotiation. I'll argue that a general scenario in language dynamics could be such that shared linguistic conventions would not emerge as attractors, but rather as metastable states.

12

Lucía Loureiro Porto
University of Palma de Majorca

"The complexity of social interactions in language competition: Agent-based models and complex networks."

This paper explores the role played by complexity, as understood in complexity theory, in the competition between languages in bilingual communities. Stemming from Abrams and Strogatz (2003) and Minett and Wang (2008), we focus on the interaction among the individuals, and our agent-based models describe the dynamics of language competition within the framework of social consensus problems. Two parameters are included in our models, namely prestige and volatility, and by changing their values we describe different sociolinguistic situations, such as language endangerment, language coexistence, and the emergence of new linguistic varieties. In addition, different topological distributions of agents are considered, within the approach of complex networks, such as regular lattice, small-world networks, and random networks (cf. Castelló et al. 2006, 2007, 2008).

References:

Abrams, D.M., Strogatz, S.H., 2003. "Modeling the dynamics of language death". *Nature* 424, 900.

Castelló, Xavier, Víctor M. Eguíluz & Maxi San Miguel. 2006. "Ordering dynamics with

two non-excluding options: Bilingualism in language competition". *New Journal of Physics* 8, 308.

Castelló, Xavier, Lucía Loureiro-Porto, Víctor M. Eguíluz & Maxi San Miguel. 2007. "The fate of bilingualism in a model of language competition". In Shingo Takahashi, David Sallach & Juliette Rouchier (eds.) *Advancing social simulation. The First World Congress*. Berlin: Springer Verlag, pp. 83-94.

Castelló, Xavier, Riita Toivonen, Víctor M. Eguíluz, Lucía Loureiro-Porto, Jari Saramäki, Kimmo Kaski & Maxi San Miguel. 2008. "Modelling language competition: Bilingualism and complex social networks". In Andrew D.M. Smith, Kenny Smith & Ramon Ferrer i Cancho (eds.) *The evolution of language. Proceedings of the 7th International Conference (EVOLANG7)*. Singapore: World Scientific Publishing Co., pp. 59-66.

Minett, James W. & William S-Y. Wang. 2008. "Modelling endangered languages: The effects of bilingualism and social structure". *Lingua* 118: 19-45.

Salikoko S. Mufwene
University of Chicago & Collegium de Lyon

"The Emergence of Complexity in Language: An Evolutionary Perspective."

An increasing number of linguists (including the present author) and other scholars especially interested in the evolution and/or the ontogenetic development of language have claimed that languages are complex adaptive systems (CAS). These have been characterized as reflecting complex dynamics of interactive agents, experiencing constant instability, and in search for equilibrium in response to changes in the ecologies of their usage. Putatively, thanks to self-organization, transitional moments of apparent stability obtain during which patterns and systems emerge, and evolutions obtain from the alternations of periods of instability and stability in seemingly unpredictable ways.
In this paper I address the following questions:

1) How many interpretations of 'complexity' apply to language(s)?
2) What and/or who are the interactive agents that produce the above characteristics?
3) From the point of view of language evolution, how did complexity emerge in language(s) and in what particular order?
4) What kind(s) of evidence support(s) the various interpretations of 'complexity' that are conceivable?

5) How does complexity in language compare with complexity in other non-linguistic phenomena?
6) What causes the “chaos” that prompts languages to reorganize themselves into new systems?

Fermin Moscoso del Prado
Dynamique du Langage, Université Lyon 2

"Linguistic complexity as optimal grammar length: An information-theoretic approach."

I will introduce a general and easily computable definition of linguistic complexity. I will show how simple information-theoretical measures can be used to compute Gell-Mann's Effective Complexity measure for language.
The measure is derived as the minimal possible size of a grammar that can generate all sentences in a reference corpus. The length of the reference corpus is then taken to its infinite limit (thus ensuring that all possible sentences in the language can be generated).
I discuss the necessary condition that must be met for the existence of a finite grammar describing a given language. Finally, through the analysis of two American English corpora (one written and the other spoken), I demonstrate the impossibility of constructing a finite grammar with full coverage of English.
I will conclude by discussing the theoretical implications of this finding.

Tom Schoenemann
Indiana University

"A complex adaptive systems approach to language and brain."

An important aspect of language complexity derives from recognizing its evolutionary origins and development. A complex adaptive systems approach is a useful framework for understanding this complexity. Complex adaptive systems are self-organizing patterns that result from repetitive adaptive interactions of many different individual agents. Language can be understood as the result of a complex adaptive system operating on the cultural and biological evolutionary levels. Languages are conventionalizations of behaviors that come to be understood to convey meaning by individuals in a social community. The brains of individuals allow them to learn these conventionalizations through repetitive interactions with other individuals, where continual reassessments are made of the patterns of association of various kinds of observable behavior (including vocalizations) and environmental events with various possible intended meanings. These conventionalizations emerge and spread through various communal populations. Learning conventionalizations also presupposes some common substrate of shared interpretation of behaviors and environmental events. This shared meaning is the legacy of human biological evolution acting on brain circuitry and a legacy of the

history of particular human communities. Some of this legacy must have predated language, but language behavior – because it is biologically adaptive – likely also induced evolutionary change in brains. Thus, multiple types of feedback in both cultural and biological systems has resulted in the phenomenon we call language, as well as modern human biology. Complexity in language derives from the multiple feedback systems underlying cultural and biological coevolution. What is known about the coevolution of brain and language will be used to illustrate some of these ideas.

Luc Steels
University of Brussels (VUB-AI Lab)
and Sony CSL, Paris

"Explaining the Origins of Complexity in Language: A case study for agreement systems."

The complexity observed in human languages (structural complexity, system complexity, size complexity, individual diversity and language diversity) can only be understood by taking a functional and evolutionary viewpoint.
The functional viewpoint insists that language is a tool for communication and hence language users expand their language systems to increase expressive power, but this must be balanced with minimizing cognitive effort, e.g. by re-using existing elements for new purposes or by limiting combinatorial search.
The evolutionary viewpoint emphasizes that a language community has to collectively find within the open-ended space of possible languages a candidate that satisfies their purposes. Nobody has a global view or telepathic insight in the brain of another person. So speakers and hearers make guesses and adjust these based on the outcome of their communications. Individual variation arises naturally and language diversity is unavoidable.

I will report on a number of computational simulations and robotic experiments of language games, in which these functional and evolutionary viewpoints have been explored.

The experiments allow us to pin down precisely the different factors contributing to the complexity of human language. I will use case studies coming from studying how (internal) agreement systems may arise.

William S-Y. Wang
Chinese University of Hong Kong

"Language and Complex Adaptive Systems."

Language is not a compact system, like the chess game that Saussure envisioned, though this view has dominated structuralist and generativist thinking over the past century. Rather, Wittgenstein's metaphor of an ancient city is much more illuminating of how languages are diffuse and heterogeneous. Each language is invariably a mixture consisting of numerous diverse pieces that have been brought together by combinations of horizontal and vertical transmission across time and space, for socio-cultural reasons having to do with its speakers, rather than with the language itself. As languages are learned, natively by the child, and as foreign languages by adults, the learner adapts to the heterogeneity by seeking generalizations among the diverse pieces. The systems learned are indeed complex. But there are great difficulties in quantifying this complexity, for reasons that will be discussed.

Some relevant references:

Cavalli-Sforza, L.L. & W.S-Y. Wang. 1986. Spatial distance and lexical replacement. *Language* 62.38-55.

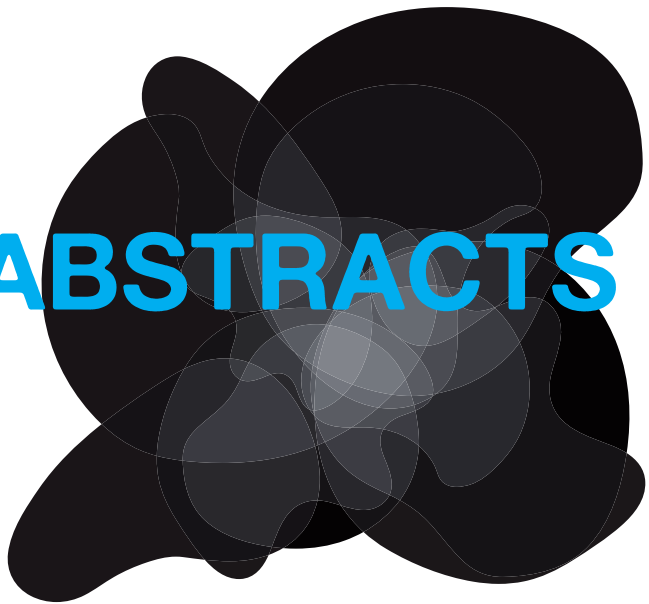
Greenberg, Joseph H. 1971. Is language like a chess game? *Language, Culture and Communication.* Pp.330-51. Stanford University Press.

Holland, John H. 1992. *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.

Minett, J.W. & W.S-Y. Wang (eds) 2005. *Language Acquisition, Change and Emergence: Essays in Evolutionary Linguistics*. Hong Kong: City University of Hong Kong Press. xii+538 pages.

Wang, W.S-Y. & C.F. Lien. 1993. Bidirectional diffusion in sound change. In *Historical Linguistics: Problems and Perspectives*, ed. by C. Jones, 345-400. Longman.

ABSTRACTS



ENS DE LYON

COLLEGIUM INSTITUT D'ÉTUDES AVANCÉES

UNIVERSITÉ LUMIÈRE LYON 2 UNIVERSITÉ DE LYON

European Research Council erc

IXXI

Dynamique Du Langage

L'ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

Région RhôneAlpes

Conception / réalisation : Joachim Bressat